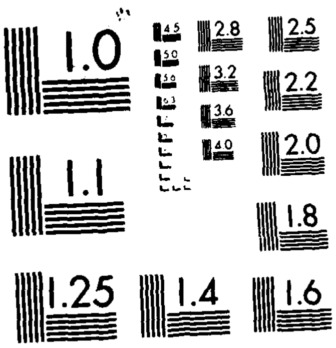
AD-A125 779 THE EFFECT OF ENVIRONMENTAL CHANGE IN SINGLE-SUBJECT EXPERIMENTS(U) DESMATICS INC STATE COLLEGE PA K C BURNS ET AL. FEB 83 TR-112-12 N00014-79-C-0128 1/1

UNCLASSIFIED F/G 12/1 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963 A

DESMATICS, INC.

P. O. Box 618
State College, PA. 16801
Phone: (814) 238-9621

Applied Research in Statistics - Mathematics - Operations Research

THE EFFECT OF ENVIRONMENTAL CHANGE
IN SINGLE-SUBJECT EXPERIMENTS

by

Kevin C. Burns
and
Dennis E. Smith

TECHNICAL REPORT NO. 112-12

February 1983

This study was supported by the Office of Naval Research
under Contract No. N00014-79-C-0128, Task No. NR 207-037

TABLE OF CONTENTS

		Page
I.	INTRODUCTION	1
II.	BACKGROUND	3
III.	STATISTICAL METHODS AND RESULTS	4
IV.	DISCUSSION	21
V.	REFERENCES	24

I. INTRODUCTION

In order to study the effects of unusual environments over a period of time, it is often necessary to analyze repeated measurements taken on a single subject. In some situations, such an analysis may be carried out under the assumption that the correlation between any two observations is the same. However, it is more usual to find that the correlations are smaller between more distant observations. While efficient inference procedures exist for the former situation, the statistical methods used in the latter case have been found to be generally inadequate.

The research presented in this report was undertaken in an attempt to develop suitable analysis techniques for single-subject experiments where the performance of the subject is abruptly affected by some environmental change (e.g., the onset of very low frequency motions which cause motion sickness). The measurements taken on the subject may be thought of as realizations of a process occurring over time. The problem considered here is that of making inferences about the mean of that process. The ultimate goal is to evaluate the statistical significance of any change in the mean which can be attributed to the environmental change.

The primary focus of this research effort has been on approximating the distribution of the standard t statistic, which is the appropriate test statistic for testing the mean of a set of observations when those observations are independent (subject to some distributional assumptions). In the situation considered here, however, autocorrelation is present. Unfortunately, the distribution of the t statistic depends upon the autocorrelations, which are difficult to estimate precisely. While some progress has been made toward understanding the distributional properties of the t statistic when autocorrelation is present, it must be concluded that this approach has not yielded satisfactory results.

II. BACKGROUND

It is well known that when observations are autocorrelated, tests based on the assumption of independence may be very misleading. Therefore, Box-Jenkins time series techniques are often employed. These methods incorporate the autocorrelation structure in the model. They require that one first identify the true model of which the observed series is a realization. Subsequent inferences depend heavily on the accuracy of this identification. Eilingstad and Westra [1] have pointed out that for short time series (less than about 50 points), model identification can be extremely difficult. The main reason for this is that model identification is dependent on the estimates of the autocorrelations. These estimates tend to be very imprecise, leading to incorrect specification of the model.

In response to the difficulty involved in model identification, Higgins [2] has studied some robustness properties of the first-order autoregressive model (AR(1)). In particular, he has considered the consequences of assuming that the correct model was an AR(1) when in fact the series was generated from a very different underlying model. Zinkgraf and Willson [4] have given similar consideration to the problem of misspecification. These articles both conclude that using an AR(1) leads to acceptable results under a wide variety of true underlying models. Therefore, this investigation focuses on making inferences about the mean of an AR(1) process. As has been shown in Higgins and elsewhere, it is straightforward to generalize the methods used to test the mean of an AR(1) to appropriate methods for testing an intervention (change in the mean) effect. Thus, only the former problem need be discussed here.

III. STATISTICAL METHODS AND RESULTS

As stated earlier, we focus attention on the AR(1) process. The model may be specified as follows:

$$Y_1 - \mu = \varepsilon_1 ,$$
$$Y_t - \mu = \rho(Y_{t-1} - \mu) + \varepsilon_t , \quad t = 2, 3, \ldots, n \text{ where}$$
$$\varepsilon_1 \sim N(0, \sigma_\varepsilon^2 /(1-\rho^2)) ,$$
$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2) \text{ for } t \geq 2 \text{ and}$$
$$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n \text{ are independent.}$$

Thus, the correlation between any two observations Y_i and Y_j is $\rho^{|i-j|}$. Also, $Y_t \sim N(\mu, \sigma^2)$ for all t where $\sigma^2 = \sigma_\varepsilon^2/(1-\rho^2)$.

The hypothesis to be tested is $H_0: \mu = 0$ vs. $H_A: \mu > 0$. A number of possible tests have been suggested for this problem. When $\rho = 0$, the appropriate test statistic is $T = \sqrt{n}\bar{Y}/s$, which follows Student's t distribution with n-1 d.f. When $\rho \neq 0$, the distribution of this statistic is unknown. From reference [2], we have that the variance of $\bar{Y}$ is given by $V(\bar{Y}) = (\sigma^2/n)\ \{1 + (2/n) \sum_{k=1}^{n} [(n-k)\rho^k]\ \}$, which may be approximated by $V(\bar{Y}) \simeq (\sigma^2/n)(\frac{1+\rho}{1-\rho})$. Using this approximation, a reasonable test statistic is given by $TC1 = \bar{Y}/(\hat{V}(\bar{Y}))^{\frac{1}{2}} = (\frac{1-\hat{\rho}}{1+\hat{\rho}})^{\frac{1}{2}} T$, where $\hat{\rho}$ denotes the estimate used for ρ. Higgins has studied the performance of this statistic for the values $n = 10, 20, 40, 80$, $\rho = .5, .8, .9$ and $\alpha = .10, .05$, where n is the sample size and α is the significance level of the test. Although this procedure does much better than that assuming independence, the empirical critical values, as calculated from 400 simulations for each (n,ρ) combination, are still not close to the critical values of the respective t distributions.

If ρ is known, the transformation $Z_t = Y_t - \rho Y_{t-1}$ may be used. Then $Z_2, Z_3, \ldots, Z_n$ are independent and identically distributed with mean

$(1-\rho)\,\mu$. Of course, in practical situations, ρ is not known. However, in practice the transformation $X_t = Y_t - \hat{\rho}Y_{t-1}$ may be used and the resulting $\{X_t\}$ treated as an independent sample. Intuitively, if $\hat{\rho}$ is a good estimate of ρ, this procedure should be close to size α. In the remainder of this report, the test statistic obtained from this procedure will be denoted TR1.

Unfortunately, the usual estimates of ρ may be very biased, especially for small samples. For $\hat{\rho} = \{\sum_{i=1}^{n-1} [(Y_i - \bar{Y})(Y_{i+1} - \bar{Y})] / \sum_{i=1}^{n} (Y_i - \bar{Y})^2\}$, Kendall and Stuart [3] give $E(\hat{\rho}) \simeq \rho - (1+3\rho)/(n-1)$. In order to obtain a less biased estimate of ρ, $\hat{\rho}^* = [(n-1)\hat{\rho} + 1]/(n-4)$ may be used. It is possible for this estimate to fall out of the interval $[-1, 1]$. In those cases, $\hat{\rho}^* = 1$ (or -1) could be used. This correction will increase the bias of $\hat{\rho}^*$ but is necessary for a sensible procedure. Both TC1 and TR1 may be modified by using $\hat{\rho}^*$ instead of $\hat{\rho}$. These modified statistics will be referred to as TC2 and TR2, respectively.

Previous studies have shown that as ρ increases, the distribution of T shows more marked departures from that of $t(n-1)$ (Student's t with $n-1$ d.f.). The differences between the two distributions are also more extreme for small sample sizes.

Preliminary analyses focused on the extremes of the design used by Higgins. That is, a 2^3 factorial design was used with $n = 10, 80$, $\rho = .5, .9$ and $\alpha = .10, .05$. For each (n,ρ) pair, 400 simulations were done. The observed critical values were found and compared to $t_\alpha(n-1)$, the upper α-point of a $t(n-1)$ distribution. Also, the empirical probability of rejecting H_0 was found when $t_\alpha(n-1)$ was used as the critical value. The procedures using $\hat{\rho}^*$ performed better than those using $\hat{\rho}$ but were still not close to size α. The results are presented in Tables 1 and 2. As can be seen from the

Table 1: Critical values for various test statistics based on 400 simulations at each (n,ρ) combination

	α	T	TC1	TC2	TR1	TR2	$t_{\alpha}(n-1)$
$n = 10,\ \rho = .5$	.10	2.51	2.07	1.71	2.28	1.55	1.383
	.05	3.16	3.10	2.65	3.24	2.40	1.833
$n = 10,\ \rho = .9$	.10	6.82	5.41	3.57	6.01	3.32	1.383
	.05	10.74	7.56	6.06	8.38	5.29	1.833
$n = 80,\ \rho = .5$	.10	2.33	1.39	1.35	1.43	1.34	1.282
	.05	2.91	1.79	1.73	1.75	1.66	1.645
$n = 80,\ \rho = .9$	.10	6.40	1.94	1.65	2.00	1.49	1.282
	.05	8.37	2.87	2.34	2.95	2.20	1.645

Table 2: Empirical significance levels when $t_{\alpha}(n-1)$ is used as the critical value

	α	T	TC1	TC2	TR1	TR2
$n = 10,\ \rho = .5$	.10	.21	.18	.14	.18	.13
	.05	.15	.12	.09	.12	.08
$n = 10,\ \rho = .9$	.10	.37	.30	.24	.32	.21
	.05	.35	.26	.17	.28	.17
$n = 80,\ \rho = .5$	.10	.25	.11	.11	.13	.11
	.05	.20	.07	.06	.06	.05
$n = 80,\ \rho = .9$	.10	.41	.20	.17	.20	.13
	.05	.38	.14	.10	.14	.08

tables, the results are not too encouraging, particularly for $n = 10$ and $\rho = .9$.

The correction used by Higgins is based on the premise that $T = wZ$, where w is some constant depending on n, ρ and α, and Z follows Student's t distribution. Therefore, $Pr(T > c) = Pr(wZ > c) = Pr(Z > c/w)$. If c is the observed critical value and $Z \sim t(n-1)$ then $w = c/t_{\alpha}(n-1)$. The values of w for the 2^3 design are given in Table 3.

As can be seen from Table 3, there is a strong effect due to ρ, as was expected. The other factors (n and α) seem to have little, if any, effect. The total variation in Table 3 can be broken up into seven components corresponding to the three main effects and their interactions. This AOV breakdown is presented in Table 4.

The AOV table shows that there is a very strong effect for ρ but no other large effects. If the three-way interaction ($n \times \alpha \times \rho$) is used as an error term, none of these other effects approaches statistical significance. Therefore, w seems to be a function of ρ alone. The next step is to try to find this function. In order to accomplish this, 1000 simulations of length 30 were done for $\rho = .5, .7, .8, .85, .9$. It was thought that, while α was not a significant factor in the 2^3 design, an effect might be found farther out in the tails of the distribution ($\alpha < .05$) . Therefore, the empirical critical values were found for $\alpha = .10, .05, .025$, and $.01$. The results of this investigation are given in Tables 5 and 6. Following the approach used by Higgins, $w = ((1+\rho)/(1-\rho))^{\frac{1}{2}}$ was used initially. The empirical critical values of $((1-\rho)/(1+\rho))^{\frac{1}{2}}T$ are presented in Table 7.

Examination of Table 7 reveals the fact the $((1-\rho)/(1+\rho))^{\frac{1}{2}}T$ does not follow a t distribution. While the empirical critical values are close to those of Student's t for $\rho = .5$, the differences become much

Table 3: Values of $w = c/t_{\alpha}(n-1)$ where c is the empirical critical value

	$\alpha = .05$		$\alpha = .10$	
	$\rho = .5$	$\rho = .9$	$\rho = .5$	$\rho = .9$
$n = 10$	1.72	5.86	1.82	4.93
$n = 80$	1.77	5.09	1.82	4.99

Table 4: AOV table for 2^3 design

Effect	df	SS
n	1	0.05
α	1	0.10
ρ	1	23.60
$n \times \alpha$	1	0.08
$n \times \rho$	1	0.07
$\alpha \times \rho$	1	0.17
$n \times \alpha \times \rho$	1	0.10

Table 5: Observed critical values based on 1000 simulations with $n = 30$ for each value of ρ. Test statistic is $T = \sqrt{n}\bar{Y}/s$.

	ρ				
α	.5	.7	.8	.85	.9
.10	2.10	3.05	3.90	4.84	6.37
.05	2.82	4.11	5.40	6.75	8.76
.025	3.41	4.97	6.86	8.31	10.89
.01	4.40	6.38	8.68	10.28	13.34

Table 6: Values of $w = c/t_\alpha(n-1)$ where c is the observed critical value

	ρ				
α	.5	.7	.8	.85	.9
.10	1.60	2.33	2.97	3.69	4.86
.05	1.66	2.42	3.18	3.97	5.15
.025	1.67	2.43	3.35	4.06	5.33
.01	1.79	2.59	3.53	4.17	5.42

Table 7: Observed critical values for $((1-\rho)/(1+\rho))^{\frac{1}{2}}T$

	ρ					
α	.5	.7	.8	.85	.9	$t_\alpha(29)$
.10	1.21	1.28	1.30	1.38	1.46	1.311
.05	1.63	1.73	1.80	1.92	2.01	1.699
.025	1.97	2.09	2.29	2.37	2.50	2.045
.01	2.54	2.68	2.89	2.93	3.06	2.462

greater as ρ increases. Therefore, a stronger correction for ρ seems to be called for. Several other functional forms were investigated. Of the functions considered, $T/(a+b/(1-\rho^2)^{\frac{1}{2}}$ was the most promising candidate. It should be noted here that for this function to be used for all positive ρ, a restriction is needed; namely, $b=a+1$. (This restriction is necessary for the function to be appropriate when $\rho=0$.) However, the function fit here is intended only to be used for $\rho>.5$. When $0<\rho<.5$, $((1-\rho)/(1+\rho))^{\frac{1}{2}}T$ fits well enough for most practical considerations.

The data used to fit the function discussed above consisted of 1000 simulations of length 30 at each value of ρ. The autocorrelations used were $\rho=.5, .7, .8, .85, .9$. The fitted function was $T/(-1.87+3.08/(1-\rho^2)^{\frac{1}{2}})$. The observed critical values are given in Table 8. If $t_\alpha(29)$ had been used as the critical value with this test statistic, the actual significance levels would have been as given in Table 9.

The results in Tables 8 and 9 are somewhat encouraging in that $T/(-1.87+3.08/(1-\rho^2)^{\frac{1}{2}})$ seems to follow Student's t distribution for this particular set of simulation data. However, as mentioned earlier, the estimates usually used for ρ are biased. Therefore, the test statistic will not follow a t distribution when the true value of ρ is not known. In order to evaluate the effect of bias in the estimate of ρ, the average values of $\hat{\rho}$ and $\hat{\rho}^*$, which will be denoted by r and r*, respectively, were computed for each set of 1000 simulations. Test statistics of the same functional form were fit to the data for each of the two estimators of ρ. While the form of the function still seems reasonable in the presence of bias, the magnitude of the coefficients is larger for more biased estimates. The fitted functions are $T/(-9.34+9.86/(1-r^2)^{\frac{1}{2}})$ and $T/(-4.33+5.08/(1-r^{*2})^{\frac{1}{2}})$ respectively. The results of this investigation are presented in Tables 10

Table 8: Observed critical values for $T/(-1.87+3.08/(1-\rho^2)^{\frac{1}{2}})$

α	ρ = .5	.7	.8	.85	.9	$t_\alpha(29)$
.10	1.25	1.25	1.19	1.22	1.23	1.311
.05	1.67	1.68	1.66	1.70	1.69	1.699
.025	2.02	2.04	2.10	2.09	2.10	2.045
.01	2.61	2.61	2.66	2.59	2.57	2.462

Table 9: Empirical significance levels when $t_\alpha(29)$ is used as the critical value

α	ρ = .5	.7	.8	.85	.9
.10	.089	.088	.084	.087	.092
.05	.047	.048	.045	.050	.049
.025	.024	.024	.028	.028	.028
.01	.014	.014	.012	.012	.013

Table 10: Observed critical values for $T/(-9.34 + 9.86/(1 - r^2)^{1/2})$

α	ρ = .5	.7	.8	.85	.9	$t_\alpha(29)$
.10	1.49	1.16	1.09	0.94	1.33	1.311
.05	1.99	1.57	1.52	1.63	1.83	1.699
.025	2.41	1.90	1.92	2.01	2.28	2.045
.01	3.12	2.43	2.44	2.48	2.79	2.462

Table 11: Empirical significance levels when $t_\alpha(29)$ is used as the critical value

α	ρ = .5	.7	.8	.85	.9
.10	.124	.075	.070	.081	.101
.05	.077	.040	.036	.042	.061
.025	.046	.020	.019	.023	.037
.01	.024	.009	.007	.010	.019

Note: The average values of $\hat{\rho}$ were:

ρ = .5	.7	.8	.85	.9
.399	.566	.645	.682	.716

Table 12: Observed critical values for $T/(-4.33+5.09/(1-r^{*2})^{\frac{1}{2}})$

			ρ			
α	.5	.7	.8	.85	.9	$t_\alpha(29)$
.10	1.42	1.21	1.12	1.18	1.29	1.311
.05	1.89	1.63	1.56	1.64	1.78	1.699
.025	2.29	1.97	1.97	2.02	2.21	2.045
.01	2.96	2.52	2.50	2.50	2.71	2.462

Table 13: Empirical significance levels when $t_\alpha(29)$ is used as the critical value

			ρ		
α	.5	.7	.8	.85	.9
.10	.114	.089	.074	.082	.098
.05	.068	.043	.041	.042	.057
.025	.036	.022	.021	.023	.033
.01	.023	.013	.011	.011	.017

Note: The average values of $\hat{\rho}^*$ were:

		ρ		
.5	.7	.8	.85	.9
.484	.670	.758	.798	.835

through 13.

It is clear from the average values obtained that $\hat{\rho}$ is extremely biased. While $\hat{\rho}^*$ is also biased, the bias is less drastic and perhaps small enough that it will not present much difficulty. From the preceding tables it is clear that the average value of $\hat{\rho}^*$ gives much better results than that of $\hat{\rho}$ in the situation being considered there.

For the data from which it was derived $T/(-1.87+3.08/(1-\rho^2)^{\frac{1}{2}})$ performs fairly well. However, it must be checked on independent data in order to ensure its general applicability. To accomplish this purpose, it was fit to the earlier 2^3 design points with $\rho = .5, .9$, $n = 10$, 80 and $\alpha = .10, .05$. Table 14 lists the observed critical values as well as the empirical significance levels when $t_\alpha(n-1)$ was used as the critical point.

From Table 14, it can be seen that while the test statistic does not perform as well on the independent data set as it does for the data used in its derivation, the results are reasonable. Unfortunately, when the statistics using r and r* were applied to the 2^3 design, the results were far from satisfactory, especially for $n = 10$. They are summarized in Tables 15 and 16.

The results in Tables 15 and 16 show the deleterious effects of bias in the estimate of the autocorrelation. Also, since this bias is greater for small samples, the results are now dependent on n as well as on ρ. Furthermore, these results only reflect the effects of bias on the test procedure. The estimates of ρ are highly variable and this variability can be expected to present further difficulties for the inference procedure. In order to obtain some indication of how the test statistics would behave when the individual estimates were used, those statistics were recalculated from the original simulation data, using $\hat{\rho}$ and $\hat{\rho}^*$ in the formulas. The

Table 14: $T/(-1.87+3.08/(1-\rho^2)^{\frac{1}{2}})$ applied to 2^3 design

(a) Critical values

	n = 10			n = 80		
	ρ			ρ		
α	.5	.9	$t_\alpha(9)$	.5	.7	$t_\alpha(79)$
.10	1.49	1.31	1.383	1.38	1.23	1.282
.05	1.87	2.07	1.833	1.72	1.61	1.645

(b) Significance levels when $t_\alpha(n-1)$ is used

	n = 10		n = 80	
	ρ		ρ	
α	.5	.9	.5	.9
.10	.115	.098	.113	.090
.05	.060	.060	.052	.045

Table 15: $T/(-9.34 + 9.86/(1 - r^2)^{\frac{1}{2}})$ applied to 2^3 design

(a) Critical values

	n = 10			n = 80		
	ρ			ρ		
α	.5	.9	$t_\alpha(9)$	.5	.9	$t_\alpha(79)$
.10	3.66	5.02	1.383	1.32	0.76	1.282
.05	4.59	7.91	1.833	1.64	0.99	1.645

(b) Significance levels when $t_\alpha(n-1)$ is used

	n = 10		n = 80	
	ρ		ρ	
α	.5	.9	.5	.9
.10	.285	.345	.105	.018
.05	.235	.297	.050	.005

Table 16: $T/(-4.33+5.09/(1-r*^2)^{1/2})$ applied to 2^3 design

(a) Critical values

	n = 10			n = 80		
	ρ			ρ		
α	.5	.9	$t_\alpha(9)$	.5	.9	$t_\alpha(79)$
.10	1.90	1.83	1.383	1.54	1.01	1.282
.05	2.36	2.89	1.833	1.92	1.33	1.645

(b) Significance levels when $t_\alpha(n-1)$ is used

	n = 10		n = 80	
	ρ		ρ	
α	.5	.9	.5	.9
.10	.152	.172	.139	.055
.05	.105	.100	.085	.023

results of those calculations are summarized in Tables 17 and 18.

Clearly, when the test statistics using $\hat{\rho}$ are calculated, their distribution is far from that of Student's t. Although using $\hat{\rho}^*$ yields better results than using $\hat{\rho}$, the distribution of the test statistic is still approximated very poorly by Student's t. The approximation is particularly bad far out in the tails of the distribution (e.g., $\alpha = .01$). This is an indication that some correction needs to be made for α. It is also clear that the correction for ρ needs to be refined. Unfortunately, further investigation failed to uncover any promising candidates for w.

Table 17: $T/(-9.34+9.86/(1-\hat{\rho}^2)^{\frac{1}{2}})$ calculated from 1000 simulations for each (n, ρ) pair

(a) Observed critical values

α	ρ = .5	.7	.8	.85	.9	$t_\alpha(29)$
.10	1.70	1.42	1.35	1.35	1.67	1.311
.05	2.66	2.21	2.16	2.24	2.54	1.699
.025	3.52	2.95	3.05	3.05	3.61	2.045
.01	4.87	5.15	5.27	5.33	6.30	2.462

(b) Significance levels when $t_\alpha(29)$ is used

α	ρ = .5	.7	.8	.85	.9
.10	.143	.103	.104	.103	.125
.05	.100	.074	.076	.080	.097
.025	.074	.060	.053	.059	.069
.01	.058	.039	.036	.043	.051

Table 18: $T/(-4.33+5.09/(1-\hat{\rho}^{*2})^{\frac{1}{2}})$ calculated from 1000 calculations for each (n,ρ) pair

(a) Observed critical values

	ρ					
α	.5	.7	.8	.85	.9	$t_\alpha(29)$
.10	1.53	1.37	1.33	1.31	1.57	1.311
.05	2.12	2.12	2.13	2.28	2.49	1.699
.025	2.78	2.70	2.90	2.98	3.64	2.045
.01	3.89	4.42	4.99	4.90	6.14	2.462

(b) Significance levels when $t_\alpha(29)$ is used

	ρ				
α	.5	.7	.8	.85	.9
.10	.128	.103	.104	.102	.121
.05	.082	.075	.076	.081	.090
.025	.054	.053	.055	.061	.064
.01	.037	.036	.035	.043	.051

IV. DISCUSSION

The focus of this report has been on hypothesis tests about the mean of a first-order autoregressive process. This type of process has been determined to be the most applicable model for the analysis of the effect of environmental change in single-subject experiments. Two types of test statistic have been considered. If the sequence of observations is denoted by $Y_1, Y_2, \ldots, Y_n$ then one type of test statistic involves estimating the autocorrelation (ρ) and then using the transformation $X_t = Y_t - \hat{\rho}Y_{t-1}$, where $\hat{\rho}$ is the estimate of ρ. The transformed observations are then treated as an independent sample.

The other type of test statistic is of the form $w(\hat{\rho})T$, where $T = \sqrt{n}\bar{Y}/s$. First, an attempt was made to determine a suitable functional form for w and then simulation data was used to estimate the coefficients. Two different estimators of ρ were considered. The first is the standard estimator while the second includes a correction intended to reduce the bias of the estimates.

The results of the previous section are not very encouraging. Test statistics which depend on transforming the data are seen to be very sensitive to the estimate of the autocorrelation. The modified estimate of ρ improves the performance of the test procedure considerably but is not sufficient for the purpose of obtaining a valid method of analysis. Not only the bias but also the variability of the estimator affect the test procedure. While further refinement of the adjustment for bias might be possible, there is no way to eliminate the variability.

Functions of the form $w(\hat{\rho})T$ attempt to compensate for the distributional properties of the estimator of ρ. Unfortunately, no suitable function has been found and perhaps none exists. It seems clear that w needs to depend on n and α (the sample size and significance level, respectively) as well as $\hat{\rho}$. That is because the sampling properties of $\hat{\rho}$ depend on n and α. At the present time, there is no clear indication as to how to proceed in finding an appropriate function $w(n,\hat{\rho},\alpha)$ so that the test statistic T/w would have an approximate t distribution. Furthermore, the existence of an adequate approximation of this type is by no means guaranteed.

Since the results presented in this report indicate that the techniques so far considered will probably not prove adequate, it is felt that a radically different approach is needed. In addition, it seems clear that such an approach will need to combine information from more than one experimental subject. Although the indiscriminate pooling of data from several individuals is not feasible in this situation, a more sophisticated approach might make such pooling possible.

Different individuals show different patterns of response to environmental change. However, it seems reasonable to assume that these individuals may be grouped into certain classes such that the individuals within each class have similar response profiles. For example, the responses from two individuals might well be easily modeled by first-order autoregressive processes with comparable autocorrelations. Even if only a few individuals may be treated as a group, the problem of testing hypotheses about the mean of the process becomes much more tractable.

At the present time, further investigation is needed in order to identify which variables best describe the response profiles of experimental subjects in environmental time-course studies. Once these variables are identified, statistical techniques such as cluster analysis may be used to separate the subjects into relatively homogeneous groups. Finally, statistical analysis techniques could be developed to make inferences about each group.

V. REFERENCES

[1] Ellingstad, V. S. and Westra, D. P., (1976). "Evaluation Methodology for Traffic Safety Programs," Human Factors, Vol. 18, pp. 313-326.

[2] Higgins, J. J. (1978). "A Robust Model for Estimating and Testing for Means in Single Subject Experiments," Human Factors, Vol. 20, pp. 717-724.

[3] Kendall, M. G. and Stuart, A. (1968). The Advanced Theory of Statistics, Vol. 3, Hafner Publishing Company, New York.

[4] Zinkgraf, S. A. and Willson, V. L. (1981). "The Use of the Box-Jenkins Approach in Causal Modeling: An Investigation of the Cost of the Misidentification of Selected Stationary Models," Time Series Analysis, North-Holland Publishing Company, New York, pp. 651-656.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 112-12	2. GOVT ACCESSION NO. AD-A125 779	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) THE EFFECT OF ENVIRONMENTAL CHANGE IN SINGLE-SUBJECT EXPERIMENTS		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Kevin C. Burns and Dennis E. Smith		8. CONTRACT OR GRANT NUMBER(s) N00014-79-C-0128
9. PERFORMING ORGANIZATION NAME AND ADDRESS Desmatics, Inc. P.O. Box 618 State College, PA 16801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 207-037
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE February 1983 13. NUMBER OF PAGES 24
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this report is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Environmental Changes
Autocorrelation
Single-Subject Experiments
AR(1) Model

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Statistical procedures for testing the mean of a first-order autoregressive model are evaluated. Two types of test statistic are considered. One involves estimating the autocorrelation and using that estimate to transform the data. The second type of test statistic is of the form T/w, where w is a function of the estimated autocorrelation and $T = \sqrt{n}\bar{Y}/s$. The usual estimation of the autocorrelation is used initially and compared to a revised estimator which provides less biased estimates. Each procedure is evaluated according to its performance on a set of simulated data.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

3-8

DT